Collegio Carlo Alberto

A General Behavioural Model of Random Choice

Paola Manzini

Marco Mariotti

No. 237

December 2011

Carlo Alberto Notebooks

www.carloalberto.org/working_papers

© 2011 by Paola Manzini and Marco Mariotti. Any opinions expressed here are those of the authors and not those of the Collegio Carlo Alberto.

A General Behavioural Model of Random Choice

Marco Mariotti*   Paola Manzini

Revised: February 2012

Abstract

We develop and study a stochastic eliminative choice procedure, called choice by random checklist (CRC). On the one hand CRC extends the idea of Tversky's Elimination by Aspects (EBA) to explain, for example, stochastically intransitive choices. On the other hand a special case of CRC is equivalent to random utility maximisation (RUM). In this way we bring within a single formal framework two leading random choice models in economics and in psychology, which are otherwise difficult to relate. Unlike EBA, which is defined recursively, we give an explicit formula for the choice probabilites in a CRC. We show that a CRC satisfies Monotonicity (extending a result of Sattath and Tversky [14]). We define new testable properties, Stochastic WARP and the Probabilistic Inequality, that are satisfied by CRC (and so, in particular, by a RUM).

JEL: D0

Keywords: Random Utility Maximization, Elimination by Aspects, Eliminative heuristics.

1 Introduction

The standard assumption of economic rationality is characterised by consistency conditions such as the Strong Axiom or Revealed Preference or Con-

*School of Economics and Finance, University of St Andrews, St. Andrews, Fife, KY16 9AL, Scotland, United Kingdom (mm210st-and.ac.uk). This research was largely carried out during a visit at Collegio Carlo Alberto (Turin).which we thank for generous financial support and hospitality.

gruence (Richter [13]). However, no pattern of data is likely to satisfy such conditions *exactly.* Data are noisy. A deterministic theory of choice could be interpreted as capturing mean behaviour, but a complete theory of choice ought to incorporate the variability around the mean as well as the mean (McFadden [10]). This has motivated *stochastic* theories of choice, the most practically influential of which - in economics - has been McFadden's ([8], [9]) multinomial logit model. In this paper we propose a stochastic theory of choice that unifies two separate strands from economics and psychology. Our theory is behavioural in the sense that it postulates a specific cognitive process that underlies choice; it is a sequential eliminative heuristics in the psychology tradition (it generalises Tversky's [15],[16] Elimination by Aspects), and yet one of its core special cases is just a reformulation of the class of models to which multinomial logit belongs, namely the class of random utility maximisers (RUM, pioneered by Luce [5], Block and Marschack [2] and Marschack [6]).

To illustrate, suppose you observe the frequencies with which an agent selects between the pairs of alternatives $\{x, y\}$, $\{y, z\}$ and $\{x, z\}$, and you find out that the agent selects x from $\{x, y\}$ half of the time, y from $\{y, z\}$ half of the time, but selects z out of $\{x, z\}$ all the time, written

$$\begin{aligned} p(x, xy) &= \frac{1}{2} = p(y, yz) \\ p(x, xz) &= 0 \end{aligned}$$

Because the agent expresses sometimes a 'preference' for x over y and sometimes for y over z, it may be thought the he should be sometimes observed to express a preference for x over z, but he doesn't. Technically, the agent suffers from stochastic intransitivity. The agent could be declared utterly irrational. A different working hypothesis is that the agent acts according to a systematic decision process that nevertheless contains some random component. There are two separate traditions in this vein. In economics and microeconometrics, the prevalent model postulates that the agent is a utility maximiser, but that utilities are stochastic: the agent's random choice is a *random utility maximiser*, in the sense that here is a set of utility functions $\mathbf{U}$ and and a probability measure π over $\mathbf{U}$ such that

$$prob\,(x \text{ is chosen from } A) = \pi\,(u \in \mathbf{U} : u(x) \geq u(y) \ \ \forall y \in A)$$

In the example, it is easily checked that the agent could be maximising the (ordinal) utility u^1 with probability $\frac{1}{2}$, utility u^2 with probability $\frac{1}{4}$ and utility

u^3 with probability $\frac{1}{4}$, with $u^1(z) > u^1(x) > u^1(y)$, $u^2(y) > u^2(z) > u^2(x)$ and $u^3(z) > u^3(y) > u^3(x)$.

A radically different tradition in psychology eschews the use of utility functions or preferences altogether, and postulates instead that choice is the result of sequential *elimination heuristics*. To explain random choice these heuristic are assumed to be stochastic. The idea of the seminal *Elimination By Aspect* (EBA) procedure is that there are properties (aspects) of the alternatives that the agent deems relevant. When choosing from a menu, the agent considers properties in succession. At each stage a property is extracted randomly, focussing only on the properties that appear in the menu. The agent eliminates the alternatives that do not possess the relevant property at each stage, until only one alternative remains. Tversky assumes that each property P has a (positive) value $U(P)$, and that at each stage each remaining property is selected with a probability equal to the ratio between its value and the total value of the remaining properties. EBA is defined implicitly trough a recursive formula:

$$prob\,(x \text{ is chosen from } A) = \frac{\sum_{P:P\cap A\neq\varnothing} U(P)\, prob\,(x \text{ is chosen from } A\cap P)}{\sum_{P:P\cap A\neq\varnothing} U(P)}$$

Note how *the probability of the order in which properties are considered in an EBA is menu-dependent*: the relative weight of property P depends on what other properties are 'available' in set A. In contrast, for a RUM the probability distribution of utilities is menu-independent. In spite of this apparent permissiveness, it is well-known that EBA cannot explain stochastic intransitivity, and therefore it cannot explain the choice probabilities of the example. In general, any pattern of choice that can be explained as an EBA can be also explained by a RUM, but strictly not viceversa.

Unlike RUM, which is a general class of models, EBA is a specific procedure. While several remarks have been made (by Tversky and others) on the relationship between RUM and EBA, it is fair to say that the picture remains unclear. We propose a general model, *Choice by Random Checklist* (CRC), that retains the two key psychological insights of EBA, namely: (1) the agent uses properties to sequentially eliminate alternatives that lack the relevant property at each stage; and (2) the order in which the properties appear is menu-dependent. However, we relax the assumption that the stochastic process that generates the order in which properties are considered is the specific one assumed for EBA.

We introduce the notion of a *mood*: fixing a set of properties, a mood is any complete sequence of properties. A mood is randomly drawn according to a - possibly menu-dependent - probability measure. The only restriction imposed on this probability measure is the following: if a mood has a higher probability in a menu A than in a menu B, then there is some alternative in A but not in B that has a property that is 'important' (it comes early in the order of properties) in that mood, whereas the alternatives in B don't have that property. For example, introducing a 'luxury' item in a menu that contains none may increase the probability that a 'luxury mood' (a mood in which luxury related properties are important) is activated. This restriction echoes the crucial feature of EBA that a property is only considered in a menu if the menu contains some alternative that has that property. The restriction is trivially satisfied when the probability of a mood does not depend on the menu.

A major difference with EBA is that, thanks to the notion of a mood, we can formulate the model in a non-recursive way. Beside making the choice probabilities more transparent, this permits us to apply our theory to arbitrary finite sets and relax the assumption that all probabilities $p(x, A \cap P)$ are well-defined (as the above definition of EBA requires), and also to derive in a simpler way certain results compared to a recursive structure.

To see how our model works in the example, suppose that there are three properties P, Q and R, and that x and z but not y have P, that only y has Q and that only z has R. We would allow the (menu-independent) probability distribution that assigns probabilities of $\frac{1}{2}$ to each of the two moods PQR and QPR, and probability zero to all other moods. These random sequences of properties, with these probabilities, are themselves not allowed by EBA.[1] But they yield the desired choice probabilities applying the EBA eliminative procedure.

It is a consequence of the results in Mandler, Manzini and Mariotti [7] (and we re-establish it here with different methods) that any RUM is a special case of CRC in which the probabilities of moods are menu-independent and vice-versa, and moreover that any utility generates a checklist and viceversa. Therefore our eliminative procedure creates a unified formal language that draws together very different models such as RUM and EBA, allowing one to

[1] We are abusing language here: strictly speaking, as noted, in EBA there are no complete sequences of properties unless all properties are possessed by some feasible alternative. Otherwise the sequences will be just partial ones.

switch from one interpretation to the other with ease. The model also shows that Tversky's insight of a menu-dependent random mechanism to generate properties, coupled with the eliminative procedure, leads to a structure that expands the explanatory ability of RUM.

We derive two testable properties of a CRC, which apply to any finite domain.

The first observable restriction on a CRC is a new probabilistic version of the Weak Axiom of Revealed Preference for deterministic choice functions, called Stochastic WARP. It says that if the entire support of the choice from A is contained in B, and x is in A, then x is chosen from A with a probability at least as great as the probability with which it is chosen from B. A notable implication of Stochastic WARP is Monotonicity, a classical property of random choice - stating that the probability of choice of an alternative from a menu cannot increase as the set expands. Sattath and Tversky [14] show that EBA satisfies Monotonicity. Thus our finding significantly generalises that result.

The second necessary property is a probabilistic choice inequality: for any two menus A and B,

$$p(x, A \cap B) \geq p(x, A) + p(x, B) - p(x, A \cup B)$$

While this property is implied almost directly by general probability laws in the case of RUM, it is trickier to establish for a CRC, in which the mood probabilities may be menu-dependent.

Beside partially characterising a CRC, the first result deepens our understanding of the behavioural content of RUMs *per se,* in view of the fact that RUM can be seen as a special case of CRC. The existing characterisations of the RUM model on a finite domain are in terms of certain complex polynomial inequalities (the Block-Marschack-Falmagne conditions) that, while useful, lack a direct behavioural interpretation. On the other hand, characterisations that are behaviourally appealing, such as those by Gul and Pesendorfer [3] or Natenzon, Gul and Pesendorfer [4], hold for special versions of RUM (on specific domains, such as that of lotteries or satisfying certain richness conditions, and with restrictions on the structure of the utility functions). For this reason, partial characterisations in terms of clearly interpretable necessary conditions on generic finite domains do fill a gap in our understanding of the restrictions on behaviour implied by the RUM (and of course of the more general CRC) model. We recall that the original discrete choice literature (the most important application of RUM) was initiated specifically to study choice problems in small finite domains.

Finally, we also give a complete characterisation of the restricted version of CRC in terms of solutions to a system of linear equations, which parallels McFadden and Richter's [12] methods for RUM.

2 The model

Let X be a universe of alternatives, and let $\mathcal{A}$ be a finite collection of nonempty finite subsets of X, called a *domain*. The probability of choice for each feasible alternative is observed for all sets in the collection $\mathcal{A}$. Formally, a *random choice* is a function $p : X \times \mathcal{A} \to [0, 1]$ such that

$$\sum_{x \in A} p(x, A) = 1$$

for all $A \in \mathcal{A}$. The number $p(x, A)$ expresses the probability that x is selected from A.

We now describe a particular procedure to generate a random choice. A *property* is a nonempty subset of X. Let $\mathcal{P}$ be a countable set of properties, let I be an initial interval of $\mathbb{N}$, and let M be a suitable label set with generic element m, called a *mood*. A *checklist* for mood m is a bijection $P^m : I \to \mathcal{P}$. In examples, we often label a checklist by simply listing the properties consecutively in the appropriate order, e.g. PQR is the label for the checklist $P^{PQR}(1) = P$, $P^{PQR}(2) = Q$ and $P^{PQR}(3) = Q$, with $\mathcal{P} = \{P, Q, R\}$.

A *random checklist* is a triple $\left(\{P^m\}_{m \in M}, M, \pi\right)$ where $\pi : \mathcal{A} \to \cdot(M)$ is a map that associates a probability measure over M with each choice set A. So $\pi^A(M')$ denotes the probability that a mood in M' is activated when the choice set is A, and we denote $\pi^A(m)$ the probability of the singleton $\{m\}$ when the choice set is A. This definition recognises that the probabilities of the various moods may be menu-induced, as occurs implicitly in the EBA model.

For all $m \in M$ and $A \in \mathcal{A}$, let us define recursively (following [7]) a sequence of 'survivor sets' $S_A(i,m)$:

$$\begin{aligned} S_A(1,m) &= A \\ S_A(i,m) &= \begin{cases} \displaystyle\bigcap_{j<i} S_A(j,m) \cap P^m(i) & \text{if } \displaystyle\bigcap_{j<i} S_A(j,m) \cap P^m(i) \neq \varnothing \\ \displaystyle\bigcap_{j<i} S_A(j,m) & \text{otherwise} \end{cases} \end{aligned}$$

We will say that x *eliminates* y (in mood m) if there is an $S_A(i,m)$ with $x \in S_A(i,m)$, $y \in \bigcap_{j<i} S_A(j,m)$ and and $y \notin S_A(i,m)$. That is, x eliminates y if there is an elimination stage i such that x possesses the relevant property for stage i whereas y lacks that property but has survived until stage i.

Definition 1 *A random choice p has the random checklist $\left(\{P^m\}_{m\in M}, M, \pi\right)$ if*

$$p(x,A) = \pi^A\left(\{m : \exists j \text{ for which } S_A(j,m) = \{x\}\}\right) \tag{1}$$

In this case we say that p is a choice by random checklist (CRC), and that $\left(\{P^m\}_{m\in M}, M, \pi\right)$ induces p.

Thus, the probability with which x is chosen from a set A is the total probability, using the probability measure π^A, of those moods m such that, using the checklist for mood m, the elimination procedure described by the sequence of survivor sets returns x as the unique selection from A after a finite number j of steps.

We will consider two (nested) restrictions on the influence of the choice set on the probability distribution on moods. The first restriction generalises a feature of EBA.

Definition 2 *The measure π is novelty-sensitive whenever $\pi^A(m) > \pi^B(m)$ for some $A, B \in \mathcal{A}$ implies that for all $y \in B$ there exists $x \in A \backslash B$ such that if $y \in P^m(j)$ then $x \in P^m(i)$ for some $i < j$.*

That is, π is novelty-sensitive if, when moving from a menu B to another menu A, the probability of a mood m increases only when some of the new alternatives have a property that none of the existing alternatives has, and that property is 'important' in mood m, in the sense of preceding the properties that are available in the old menu A. The agent becomes more likely to have a certain mood only if there is some newly available alternative that 'triggers' that mood because that alternative possesses an important property in the mood.

An extreme example of a similar feature is the EBA model, in which the agent restricts attention only to properties that some of the alternatives in the menu have. A property moves from being considered with zero probability to being considered with positive probability only if some new alternative appears that possesses that property, whereas previously no alternative had that property.

A simpler restriction is the following.

Definition 3 *The measure π is menu-independent if $\pi^A(m) = \pi^B(m)$ for all $A, B \in \mathcal{A}$ and $m \in M$.*

Observe that a menu-independent π is trivially also novelty-sensitive.

3 Characterisation

Given a checklist P^m, write

$$P_x^m(i) = \begin{cases} 1 & \text{if} \quad x \in P^m(i) \\ 0 & \text{if} \quad x \notin P^m(i) \end{cases}$$

Also, define the function $f^m : X \times X \times I \to \{0, 1\}$ by:

$$f^m(x, y, i) = \prod_{j=1}^{i-1} \left(1 - \left|P_y^m(j) - P_x^m(j)\right|\right)$$

where we use the notational convention that $\prod_{j=1}^{0} \alpha_j = 1$ for any argument α_j. The function f^m is a 'tie prior to i indicator function' in the sense that $f^m(x, y, i) = 1$ if for all $j < i$ either $x, y \in P^m(j)$ or $x, y \notin P^m(j)$), and $f^m(x, y, i) = 0$ otherwise.

Finally, in the sequel it is convenient to use the following shorthand

$$x = c(A, m) \Leftrightarrow \exists j : S_A(j, m) = \{x\}$$

to denote the situation in which an alternative x is the unique survivor in a set A after applying the elimination procedure in mood m.

With this notation established, in our first result we derive an explicit, non-recursive numerical formula for the probability of choice with a random checklist.

Proposition 4 *(Non-recursive formulation) If p has the random checklist $\left(\{P^m\}_{m\in M}, M, \pi\right)$ then for all $(x, A) \in X \times \mathcal{A}$ with $x \in A$*

$$p(x, A) = \sum_{m\in M} \pi^A(m) \prod_{y\in A\setminus\{x\}} \sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) \tag{2}$$

Proof. Fix m for the moment. We have $P_x^m(i)\left(1 - P_y^m(i)\right) > 0$ only for those i for which $P_x^m(i) > P_y^m(i)$, that is, $x \in P^m(i)$ and $y \notin P^m(i)$.

Therefore,

$$\sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) = 1$$

if and only if there exists i such that (1) $x \in P^m(i)$ and $y \notin P^m(i)$ and (2) there exists no $j < i$ such that either $x \in P^m(j)$ and $y \notin P^m(j)$ or $x \notin P^m(j)$ and $y \in P^m(j)$; and

$$\sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f(x, y, i) = 0$$

otherwise. Therefore if

$$\sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) = 1$$

for all y in $A\setminus\{x\}$, then $x \in S_A(j, m)$ for all j, since by the previous argument no y can eliminate x at any stage. This means (given that p has the random checklist $\left(\{P^m\}_{m\in M}, M, \pi\right)$) that $x = c(A, m)$. We now turn to the converse of this statement. If $x = c(A, m)$, then unless $P_x(i) = P_y(i)$ for all i, we can define the largest i for which $f^m(x, y, i) = 1$, denoted $i(x, y)$. We have by construction $P_x^m(i(x, y))\left(1 - P_y^m(i(x, y))\right) = 1$. If $P_x^m(i(x, y)) = 0$ and $P_y^m(i(x, y)) = 1$, then $x \notin P^m(i(x, y))$ and $y \in P^m(i(x, y))$. Then as x is not eliminated by any other alternative it is eliminated by y, a contradiction. We conclude that $P_x^m(i(x, y)) = 1$ and $P_y^m(i(x, y)) = 0$ and thus

$$\sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) = 1$$

Supposing on the other hand that $P_x^m(i) = P_y^m(i)$ for all i, then since $y \neq c(A, m)$ there exists a z that eliminates y and that z must also eliminate x, a contradiction.

So we have shown that

$$\begin{aligned} x &= c(A, m) \\ &\Leftrightarrow \sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) = 1 \ \forall y \in A \backslash \{x\} \\ &\Leftrightarrow \prod_{y \in A \backslash \{x\}} \sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) = 1 \end{aligned}$$

Thus the probability $\pi^A(m)$ in the formula for $p(x, A)$ in the statement is multiplied by one if and only if $x = c(A, m)$, and by zero otherwise. We conclude that formulae 2 and 1 coincide. ∎

This result makes explicit the connection between properties and probabilities of choice, which remains hidden in EBA. While the arguments in the sums in i of the formula of proposition 4 are simple indicator functions, a key feature is that those sums can be combined *multiplicatively*. This means in particular that in order to be selected in a mood m by the application of the corresponding checklist, an alternative must *itself* eliminate all unchosen alternatives.

3.1 Testable restrictions

In the statement of the axioms below, whenever we use an expression of the type $p(x, A)$ we take it as understood that it is well-defined, that is $A \in \mathcal{A}$.

We introduce a restriction on a random choice.

Probabilistic Inequality. $p(x, A \cap B) \geq p(x, A) + p(x, B) - p(x, A \cup B)$.

To interpret the Probabilistic Inequality, consider that the right hand side expresses the difference between the probability that an alternative x eliminates all the elements of two menus separately and the probability that it eliminates them when they are present in the same menu. The Probabilistic Inequality poses an upper bound on this difference, equal to the probability that x eliminates the alternatives that are common to both menus.

Theorem 5 *A CRC with a novelty-sensitive π satisfies the Probabilistic Inequality.*

Proof. Let p have a CRC. For any $C \in \mathcal{A}$ define

$$M_x(C) = \left\{ m \in M : \prod_{z \in C \setminus \{x\}} \sum_i \left(P_x^m(i)\left(1 - P_z^m(i)\right)\right) f^m(x, z, i) = 1 \right\}$$

and recall that $p(x, C) = \pi^C(M_x(C)) = \sum_{m \in M_x(C)} \pi^C(m)$.

First we derive an expression for $p(x, A) + p(x, B) - p(x, A \cup B)$ in terms of the measure π. It is easy to check that formula 4 implies that

$$M_x(A) = M_x(A \cup B) + M_x(A) \setminus M_x(A \cup B)$$

Similarly,

$$M_x(B) = M_x(A \cup B) + M_x(B) \setminus M_x(A \cup B)$$

Thus

$$\begin{aligned} p(x, A) &= \pi^A(M_x(A)) \\ &= \pi^A(M_x(A \cup B)) + \pi^A(M_x(A) \setminus M_x(A \cup B)) \end{aligned}$$

and

$$\begin{aligned} p(x, B) &= \pi^B(M_x(B)) \\ &= \pi^B(M_x(A \cup B)) + \pi^B(M_x(A) \setminus M_x(A \cup B)) \end{aligned}$$

However

$$\pi^A(M_x(A \cup B)) = \pi^{A \cup B}(M_x(A \cup B))$$

since if for some $m \in M_x(A \cup B)$ we had $\pi^A(m) > \pi^{A \cup B}(m)$ then because π is novelty-sensitive there should exist a $y \in A \setminus (A \cup B)$, a contradiction; and if $\pi^A(m) < \pi^{A \cup B}(m)$ then the facts that π is novelty-sensitive would imply that $m \notin M_x(A \cup B)$, again a contradiction.

By a symmetric reasoning we establish that

$$\pi^B(M_x(A \cup B)) = \pi^{A \cup B}(M_x(A \cup B))$$

We can thus write

$$\begin{aligned} & p(x, A) + p(x, B) - p(x, A \cup B) \\ = \; & \left(\pi^{A \cup B}(M_x(A \cup B)) + \pi^A(M_x(A) \setminus M_x(A \cup B))\right) + \\ & \left(\pi^{A \cup B}(M_x(A \cup B)) + \pi^B(M_x(B) \setminus M_x(A \cup B))\right) - \\ & \pi^{A \cup B}(M_x(A \cup B)) \\ = \; & \pi^{A \cup B}(M_x(A \cup B)) + \pi^A(M_x(A) \setminus M_x(A \cup B)) + \pi^B(M_x(B) \setminus M_x(A \cup B)) \end{aligned}$$

Next, we establish a lower bound for $\pi^{A\cap B}(M_x(A\cap B))$. By the definition of $M_x(.)$, if $m \in M_x(A\cup B)$ then $m \in M_x(A\cap B)$. Similarly, if $m \in M_x(A)\backslash M_x(A\cup B)$ then $m \in M_x(A\cap B)$, and if $m \in M_x(B)\backslash M_x(A\cup B)$ then $m \in M_x(A\cap B)$. Thus it follows that

$$\begin{aligned}\pi^{A\cap B}(M_x(A\cap B)) \geq\ & \pi^{A\cap B}(M_x(A\cup B)) + \pi^{A\cap B}(M_x(A)\backslash M_x(A\cup B)) + \\ & \pi^{A\cap B}(M_x(B)\backslash M_x(A\cup B))\end{aligned}$$

Now note that if $\pi^{A\cup B}(m) > \pi^{A\cap B}(m)$ then, given that p has a novelty-sensitive π, $m \notin M_x(A\cup B)$, and so $\pi^{A\cup B}(m) \leq \pi^{A\cap B}(m)$ for all $m \in M_x(A\cup B)$. Similarly, if $\pi^{A}(m) > \pi^{A\cap B}(m)$ (resp. $\pi^{A}(m) > \pi^{A\cap B}(m)$) then $m \notin M_x(A)$ (resp. $m \notin M_x(B)$), and thus $m \notin M_x(A)\backslash M_x(A\cup B)$ (resp. $m \notin M_x(B)\backslash M_x(A\cup B)$). This means that

$$\pi^{A}(M_x(A)\backslash M_x(A\cup B)) \leq \pi^{A\cap B}(M_x(A)\backslash M_x(A\cup B))$$

for all $m \in M_x(A)\backslash M_x(A\cup B)$ and that

$$\pi^{B}(M_x(B)\backslash M_x(A\cup B)) \leq \pi^{A\cap B}(M_x(B)\backslash M_x(A\cup B))$$

for all $m \in M_x(B)\backslash M_x(A\cup B)$.

Therefore the previously displayed lower bound for $\pi^{A\cap B}(M_x(A\cap B))$ implies that

$$\begin{aligned}\pi^{A\cap B}(M_x(A\cap B)) \geq\ & \pi^{A\cup B}(M_x(A\cup B)) + \pi^{A}(M_x(A)\backslash M_x(A\cup B)) + \\ & \pi^{B}(M_x(B)\backslash M_x(A\cup B))\end{aligned}$$

and then, using the previously found expression for $p(x,A) + p(x,B) - p(x,A\cup B)$,

$$p(x,A\cap B) \geq p(x,A) + p(x,B) - p(x,A\cup B)$$

as desired. ■

We now move to a new restriction on random choice:

Stochastic WARP: $\{y : p(y,A) > 0\} \subset B$ and $x \in A \Rightarrow p(x,A) \geq p(x,B)$.

When $p(x,B) = 0$, the conclusion of property follows trivially. The substance of the axiom lies in the case when $p(x,B) > 0$. Then Stochastic WARP says that if all alternatives chosen with positive probability in A are available in B, and x is chosen with positive probability from B, then it must be chosen with at least as great a probability from A.

For a deterministic, single valued choice function - identified with a random choice with probabilities of choice taking on values of either zero or one - Stochastic WARP reduces to standard WARP: if y is selected from A, that is $p(y,A) = 1$, and $x, y \in A \cap B$, then $p(x,B) \leq p(x,A) = 0$, that is x is not selected from B.

Theorem 6 *A CRC with novelty-sensitive π satisfies Stochastic WARP.*

Proof: Let p be induced by the random checklist $\left(\{P^m\}_{m \in M}, M, \pi\right)$. Let A,$B \in \mathcal{A}$ be such that $\{y : p(y,A) > 0\} \subset B$, and let $x \in A$ be such that $p(x,B) > 0$. Let $M^*_x(B)$ be the set of all $m \in M$ such that $\pi^B(m) > 0$ and

$$\prod_{z \in B \setminus \{x\}} \sum_i \left(P^m_x(i)\left(1 - P^m_z(i)\right)\right) f^m(x,z,i) = 1$$

From formula 2 $M^*_x(B)$ is nonempty and $p(x,B) = \pi^B(M^*_x(B))$. Since $\{y : p(y,A) > 0\} \subset B$, then for all $m \in M^*_x(B)$ we have

$$\prod_{z \in \{y : p(y,A) > 0\}} \sum_i \left(P^m_x(i)\left(1 - P^m_z(i)\right)\right) f^m(x,z,i) = 1 \tag{3}$$

Next, take $w \in A \setminus B$. Since by assumption $p(w,A) = 0$, for any $m \in M^*_x(B)$ either there is some $z \in \{y : p(y,A) > 0\}$ that eliminates w from A in mood m, or there is no such z but $\pi^A(m) = 0$. In the latter case we have $\pi^B(m) > \pi^A(m)$ and therefore since since π is novelty-sensitive there exists $y \in B \setminus A$ and i such that $P^m_y(i) = 1$ and $i < j$ for all j for which $P^m_x(j) = 1$, so that $m \notin M^*_x(B)$, a contradiction. Therefore the first possibility holds and then we have

$$\sum_i \left(P^m_z(i)\left(1 - P^m_w(i)\right)\right) f^m(z,w,i) = 1$$

But by equation 3 also

$$\sum_i \left(P^m_x(i)\left(1 - P^m_z(i)\right)\right) f^m(x,z,i) = 1$$

This last equation tells us that there is a j such that $P_z^m(j) = 1$, $P_w^m(j) = 0$ and $P_z^m(k) = P_w^m(k)$ for all $k < j$. The previous equation says that there is a j' such that $P_x^m(j') = 1$, $P_z^m(j') = 0$ and $P_x^m(k) = P_z^m(k)$ for all $k < j'$. If $j' < j$ we have that $P_x^m(j') = 1$, $P_z^{m^*}(j') = P_w^m(k) = 0$ and $P_x^m(k) = P_w^m(k)$ for all $k < j'$. If $j < j'$ we have that $P_x^m(j') = 1 = P_z^m(j')$, $P_w^m(k) = 0$ and $P_x^m(k) = P_w^m(k)$ for all $k < j$. In both cases,

$$\sum_i \left(P_x^m(i)\left(1 - P_w^m(i)\right)\right) f^m(x, w, i) = 1$$

(it cannot be $j = j'$ since $P_z^m(j') = 0$ and $P_z^m(j) = 1$).

Therefore,

$$\prod_{z \in A \setminus \{x\}} \sum_i \left(P_x^m(i)\left(1 - P_z^m(i)\right)\right) f^m(x, z, i) = 1$$

for all $m \in M_x^*(B)$. If $\pi^A(m) \geq \pi^B(m)$ for all $m \in M_x^*(B)$ the desired conclusion $p(x, A) \geq p(x, B)$ follows immediately. But if $\pi^A(m) < \pi^B(m)$ for some m, then since π is novelty-sensitive we have $m \notin M_x^*(B)$ by same argument used previously, which concludes the proof. ∎

A weaker classical property of random choice, a natural stochastic adaptation of the Chernoff axiom, is the following.

Monotonicity: $x \in A \subset B \Rightarrow p(x, A) \geq p(x, B)$.[2]

Monotonicity is a sufficiently strong property that, on specific domains, it is the core axiom in characterising (an expected utility version of) RUM (Gul and Pesendorfer [3]). Stochastic WARP strengthens Monotonicity: Suppose Stochastic WARP holds, and let $x \in A \subset B$. Then since $\{y : p(y, A) > 0\} \subset B$, by Stochastic WARP $p(x, A) \geq p(x, B)$ if $p(x, B) > 0$, so that Monotonicity is verified. Therefore we have immediately the following

Corollary 7 *A CRC with a novelty-sensitive π satisfies Monotonicity.*

This result clarifies theorem 2 of Sattath and Tversky [14], the first half of which states that an EBA satisfies Monotonicity. Theorem 6 and Corollary

[2] This property is also called 'regularity' in the literature. We follow the terminology of Gul, Natenzon and Pesendorfer [4].

7 show that it is not just the special probabilistic structure of an EBA that yields this property. A stronger testable restriction holds for a far larger class of EBA-like eliminative procedures. But it also holds, for example, in the case in which the probability of a mood does not depend at all on the menu - and thus one in which the choice procedure will look very unlike that of an EBA. Interestingly, the Monotonicity result could also be proved directly as a rather trivial implication of our model, thus considerably simplifying the inductive argument in [14]. This is a benefit of the non-recursive formulation of our eliminative procedure.

From an opposite perspective, it is well-known that a RUM satisfies Monotonicity (Block and Marschack [2]). Our result clarifies that it is neither the menu-independence of the probability measure on utilities, nor the utility structure itself (as opposed to an eliminative heuristics) that yield Monotonicity.

4 Linear system characterisation for a menu-independent π

When π is menu-independent we can characterise a CRC as the solution of a linear system of equations. In this section we assume that π is menu-independent and drop the superscript from the notation $\pi^A(M)$.

We introduce an auxiliary concept. A *simple checklist* is a checklist with as many properties as the alternatives, and such that each property, called a *simple property*, includes only one alternative (and thus each alternative has a property). A *simple random checklist* is a random checklist using only simple checklists. To illustrate, consider again the stochastic intransitivity example of the introduction:

$$\begin{aligned} p(x, xy) &= \frac{1}{2} = p(y, yz) \\ p(x, xz) &= 0 \end{aligned}$$

The random checklist we discussed in the introduction to explain p is not simple, but a simple random checklist that induces p is as follows:

	P	Q	R
x	1	0	0
y	0	1	0
z	0	0	1

with $\pi(RPQ) = \frac{1}{2}$ and $\pi(QRP) = \frac{1}{2}$.

A simple checklist should not be seen as a realistic description of a decision process but just as a useful theoretical tool. Given a random checklist $\left(\{P^m\}_{m\in M}, M, \pi\right)$, for each mood m, for each (x, A), write

$$g_m(x, A) = \prod_{y\in A\setminus\{x\}} \sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i)$$

which in the case of a simple random checklist simplifies to

$$g_m(x, A) = \prod_{y\in A\setminus\{x\}} \sum_i \left(P_x^m(i)\right) \prod_{j<i} \left(1 - P_y^m(j)\right)$$

In other words, as the proof of proposition 4 has shown, $g_m(x, A) = 1$ if $x = c(A, m)$ and $g_m(x, A) = 0$ otherwise. Observe that if p has a simple checklist $\left(\{P^m\}_{m\in M}, M, \pi\right)$ then it also has any other simple checklist obtained from the original one by permuting the moods and the probabilities $\pi(m)$. So, fixing a set M with simple properties and with the $g_m(x, A)$ referring to that fixed set, by definition, a random choice p has a simple random checklist if and only if there are probabilities $\pi(m)$ such that

$$p(x, A) = \sum_m \pi(m)\, g_m(x, A)$$

It is convenient to write these equations in matrix form. Enumerate the pairs $(x, A) \in X\times\mathcal{A}$ for which $x \in A$ with a bijection $b : X \times \Sigma \to \mathbb{N}$ and let $\hat{p}$ be the column vector with the entries $p(x, A)$ in the enumeration order, i.e. $\hat{p}_i = p\left(b^{-1}(i)\right)$. Enumerate the moods with $d : M \to \mathbb{N}$. Correspondingly, let G be the $|\,(x, A) : A \in \mathcal{A}\&x \in A| \times |M|$ matrix with ij entry equal to $g_{d^{-1}(j)}\left(b^{-1}(i)\right)$ (so each column corresponds to a mood and each row to an alternative-set pair). Then

Proposition 8 *A random choice p has a random checklist if and only if the linear system*

$$\hat{p} = G\pi$$

has a solution in π such that $\sum_m \pi(m) = 1$ and $\pi(m) \geq 0$ for all m.

The statement is obviously true by construction if we replace 'random checklist' with 'simple random checklist', and a simple random checklist is a checklist. So we only need to prove the following

Lemma 9 *A random choice p has a random checklist only if it has a simple random checklist.*

Proof: Suppose that p has a random checklist $\left(\{P^m\}_{m\in M}, M, \pi\right)$. For all $(x, A) \in X \times \mathcal{A}$ with $x \in A$ define $g_m(x, A)$ as in the text. Construct a simple random checklist $\left(\left\{\hat{P}^m\right\}_{m\in\hat{M}}, \hat{M}, \pi\right)$ with $\hat{M} = M$, $\hat{\pi} = \pi$, and checklist $\hat{P}^m$ for mood m satisfying the following conditions:

$$\begin{aligned} \forall (x, A) &\in X \times \mathcal{A} \text{ such that } g_m(x, A) = 1, \forall y \in A : \\ \hat{P}^m_x(i) &= 0 \Rightarrow \hat{P}^m_y(i) = 0 \end{aligned} \tag{4}$$

(i.e. select $\{x\}$ as a property preceding all other properties $\{y\}$, $y \in A\backslash\{x\}$, in mood m) . This ensures that in mood m any $y \in A\backslash\{x\}$ is eliminated by x and that no such y can eliminate x, and therefore that the simple random checklist $\left(\left\{\hat{P}^m\right\}_{m\in\hat{M}}, \hat{M}, \pi\right)$ induces the same random choice as $\left(\{P^m\}_{m\in M}, M, \pi\right)$. We need to check that this construction is possible, that is that the conditions 4 do not imply both $\hat{P}^m_x(i) = 0 \Rightarrow \hat{P}^m_y(i) = 0$ and $\hat{P}^m_y(i) = 0 \Rightarrow \hat{P}^m_x(i) = 0$ for some $m \in M$ and $x, y \in X$. Suppose to the contrary that there exist two sets A and B, a mood m, and two distinct alternatives x,$y \in A \cap B$ for which the conditions 4 applied to A requires $\hat{P}^m_x(i) = 0 \Rightarrow \hat{P}^m_y(i) = 0$ and the conditions 4 applied to B require $\hat{P}^m_y(i) = 0 \Rightarrow \hat{P}^m_x(i) = 0$. This can only happen if $g_m(x, A) = 1 = g_m(y, B)$, that is, from the definitions, if

$$\begin{aligned} \prod_{z\in A\backslash\{x\}} \sum_i \left(P^m_x(i)\left(1 - P^m_z(i)\right)\right) f^m(x, z, i) &= 1 \\ \prod_{z\in B\backslash\{y\}} \sum_i \left(P^m_y(i)\left(1 - P^m_z(i)\right)\right) f^m(y, z, i) &= 1 \end{aligned}$$

and given that $y \in A\backslash\{x\}$ and $x \in B\backslash\{y\}$, in particular

$$\begin{aligned} \sum_i \left(P^m_x(i)\left(1 - P^m_y(i)\right)\right) f^m(x, y, i) &= 1 \\ \sum_i \left(P^m_y(i)\left(1 - P^m_x(i)\right)\right) f^m(y, x, i) &= 1 \end{aligned}$$

This is a contradiction, since the top equality means that there exists i such that (1) $x \in P^m(i)$ and $y \notin P^m(i)$ and (2) there exists no $j < i$

such that either $x \in P^m(j)$ and $y \notin P^m(j)$ or $x \notin P^m(j)$ and $y \in P^m(j)$, and therefore the bottom equality cannot hold. ■

The characterisation of Proposition 8 gives a useful practical way to check whether a random choice has a random checklist or not.

To illustrate, we work out the three-alternative case. Let $X = \{x, y, z\}$. The properties in a canonical checklist are $\{x\}$, $\{y\}$, and $\{z\}$. We have six moods, m_k with $k = 1, ..., 6$, corresponding to the six possible orders in which the properties appear.

A random choice p has a simple random checklist if and only if each $p(x, A)$ can be written as

$$p(x, A) = \sum_{k=1}^{6} \pi(k)\, g_k(x, A)$$

In our case there are five free parameters $p(x, A)$: $p(x, xy)$, $p(x, xz)$, $p(y, yz)$, $p(x, xyz)$ and $p(y, xyz)$. The associated linear system is

$$\begin{pmatrix} p(x, xy) \\ p(x, xz) \\ p(y, yz) \\ p(x, xyz) \\ p(y, xyz) \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \begin{pmatrix} \pi(1) \\ \pi(2) \\ \pi(3) \\ \pi(4) \\ \pi(5) \\ \pi(6) \end{pmatrix}$$

The solutions to the system which in addition satisfy the constraint that $\sum_{k=1}^{6} \pi(k) = 1$, are given (Scientific Workplace) by:

$$\pi = \begin{pmatrix} p(y, yz) - p(y, xyz) \\ p(x, xyz) - p(y, yz) + p(y, xyz) \\ p(x, xz) - p(x, xyz) \\ p(x, xyz) - p(x, xz) + p(y, xyz) \\ p(x, xy) - p(x, xyz) \\ 1 - p(y, xyz) - p(x, xy) \end{pmatrix}$$

So any random choice p for which $\pi(k) \geq 0$ has a random checklist. As can be seen, these conditions for the first, third and fifth equation directly express some necessary monotonicity conditions for x and y. The conditions for the second and fourth equations are also monotonicity conditions, for z. E.g. the second condition says that

$$\begin{aligned} p(x, xyz) + p(y, xyz) &\geq p(y, yz) \Leftrightarrow \\ 1 - p(x, xyz) - p(y, xyz) &\leq 1 - p(y, yz) \Leftrightarrow \\ p(z, xyz) &\leq p(z, zy) \end{aligned}$$

Similarly, the fourth condition says that

$$p(z, xyz) \leq p(z, zx)$$

And the last condition is the final monotonicity condition for y: it says that

$$\begin{aligned} 1 - p(x, xy) &\geq p(y, xyz) \Leftrightarrow \\ p(y, xy) &\geq p(y, xyz) \end{aligned}$$

We have thus shown:

Corollary 10 *If $|X| = 3$, then a random choice with menu-independent π has a random checklist if and only if it satisfies Monotonicity.*

The characterisation of corollary 10 is the same as the corresponding characterisations for RUM given by Block and Marschack [2] (theorems 5.1 and 5.3). So we now turn to making explicit the connection between random checklists and random utility maximisers.

5 Relationship with Random Utility Maximisers

A random choice p is a *random utility maximiser (RUM)* if there exists a set $\mathbf{U}$ of utility functions $u : X \to \mathbb{R}$ and a probability measure λ over $\mathbf{U}$ such that $p(x, A) = \lambda(\{u \in \mathbf{U} : u(x) \geq u(y) \ \ \forall y \in A\})$.

Proposition 11 *Let p have the random checklist $\left(\{P^m\}_{m \in M}, M, \pi\right)$ where π is menu-independent. Then, defining for each $m \in M$*

$$u^m(x) = \sum_{y \in X} \sum_{i} \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i)$$

we have

$$p(x, A) = \pi(m : u^m(x) \geq u^m(y) \ \ \forall y \in A)$$

Hence, in particular, a CRC with a menu-independent π is a RUM.

Proof: we show that for all $m \in M$

$$x = c(A, m) \Leftrightarrow x = \arg\max_{a \in A} u^m(a)$$

which proves the result in view of 1.

Fix $m \in M$ and let $x = c(A, m)$ for some $A \in \Sigma$. For all $z \in A \setminus \{x\}$ and all $y \in X$,

$$\begin{aligned} \sum_i \left(P_z^m(i)\left(1 - P_y^m(i)\right)\right) f^m(z, y, i) &= 1 \Rightarrow \\ \sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) &= 1 \end{aligned} \tag{5}$$

For suppose to the contrary that the top equality of 5 holds but

$$\sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) = 0$$

If $P_x^m(i) = P_y^m(i)$ for all i we have a contradiction between

$$\sum_i \left(P_z^m(i)\left(1 - P_x^m(i)\right)\right) f^m(z, x, i) = 0$$

(which follows from $x = c(A, m)$ and $z \in A$) and the top equality of 5, once we replace $P_y^m(i)$ with $P_x^m(i)$. Then $P_x^m(j) \neq P_y^m(j)$ for some j, so that there must be a largest i for which $f^m(x, y, i) = 1$. Denote this i by $i(x, y)$, and observe that $P_x(i(x, y)) = 0$ and $P_y(i(x, y)) = 1$. Analogously, denote $i(z, y)$ as the largest i for which $f^m(z, y, i) = 1$. If $i(z, y) < i(x, y)$ then we have a contradiction with $x = c(A, m)$ and $z \in A$, as $P_x(j) = P_y(j)$ for all $j < i(x, y)$ and so the first property that distinguishes between z and y must also distinguish between z and x, i.e. $z \in P_{i(z,y)}$ and $y \notin P_{i(z,y)} \Rightarrow x \notin P_{i(z,y)}$. The same contradiction is reached if $i(z, y) \geq i(x, y)$, since $P_y(j) = P_z(j)$ for all $j < i(z, y)$.

This argument implies that

$$\begin{aligned} & \sum_{y \in X} \sum_i \left(P_x^m(i)\left(1 - P_y^m(i)\right)\right) f^m(x, y, i) \\ \geq & \sum_{y \in X} \sum_i \left(P_z^m(i)\left(1 - P_y^m(i)\right)\right) f^m(z, y, i) \end{aligned}$$

for all $z \in A \setminus \{x\}$. To see that the inequality holds strictly, recall from the proof of proposition 4 that

$$x = c(A, m) \Leftrightarrow \sum_i \left(P_x^m(i) \left(1 - P_y^m(i)\right)\right) f^m(x, y, i) = 1 \; \forall y \in A \setminus \{x\}$$

so that

$$\begin{aligned} \sum_i \left(P_z^m(i) \left(1 - P_x^m(i)\right)\right) f^m(z, x, i) &= 0 \\ \sum_i \left(P_x^m(i) \left(1 - P_z^m(i)\right)\right) f^m(z, x, i) &= 1 \end{aligned}$$

for all $z \in A \setminus \{x\}$. ■[3]

The converse of this result is easily proved by making use of lemma 9. Any simple checklist obviously identifies a strict ranking of alternatives and therefore a utility function. So the first part of the following result is immediate:

Proposition 12 *A RUM is a CRC with menu-independent π. There are CRC that are not RUM.*

The relationship between RUM and CRC with menu-independent π is also a consequence of the general results of [7], proved with less elementary methods[4]. To see that the second part of the proposition holds, consider the following example.

Let $\mathcal{A} = \{A, B\}$ with $A = \{x, y, v, w\}$ and $B = \{x, y, w, z\}$. Consider simple checklists with two moods:

$$\begin{aligned} m_1 &= \{v\}\{w\}\{x\}\{y\}\{z\} \\ m_2 &= \{z\}\{y\}\{x\}\{w\}\{v\} \end{aligned}$$

[3] The same result could also be obtained by applying the slightly more complex arguments in MMM, which shows that every deterministic choice function that has a checklist maximises a utility function, by invoking the fact a countable sequence of zeros and ones can be seen as the binary expansion of a real number, and showing that that number can serve as a utility representation for the underlying alternative.

[4] As they deal with non-finite menus, they use checklists with a countable number of properties, and exploit the fact that the zero-one sequences associated with a an alternative and a checklist can be interpreted as the binary expansion of a linear number.

Let

$$\begin{aligned}
\pi^A(m_1) &= \frac{1}{2} > \pi^B(m_1) = \frac{1}{3} \\
\pi^A(m_2) &= \frac{1}{2} < \pi^B(m_2) = \frac{2}{3}
\end{aligned}$$

so that

$$\begin{aligned}
p(v, A) &= p(y, A) = \frac{1}{2} \\
p(w, A) &= p(x, A) = 0 \\
p(w, B) &= \frac{1}{3} \\
p(z, B) &= \frac{2}{3} \\
p(x, B) &= p(y, B) = 0
\end{aligned}$$

This random choice does not have a CRC with menu-independent π. To see this, suppose on the contrary that it does. In order for both $p(y, A) = \frac{1}{2}$ and $p(y, B) = 0$ to hold, it must be that $\{z\} <_m \{y\}$ for all $m \in M_y(A)$, and that $\pi(M_y(A)) = \frac{1}{2}$. On the other hand, since $p(z, B) = \frac{2}{3} > \frac{1}{2}$, it must be that $M_z(B) \setminus M_y(A) \neq \varnothing$, with $\pi(M_z(B) \setminus M_y(A)) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$. For all $m \in M_z(B) \setminus M_y(A)$ it must be that $\{z\} <_m \{s\} <_m \{y\}$ for some $s \in \{w, x\}$. In either case a contradiction is generated, since if $\{s\} <_m \{t\}$ for all $t \in B \setminus \{z\}$, with $s \in \{w, x\}$, this would imply $p(s, A) > 0$, contradicting the assumption $p(x, A) = 0 = p(w, A)$. Therefore, by the first part of proposition 12, p is not a RUM.

An important aspect of a random checklist with menu-independent π when viewed as a RUM is that it still naturally incorporates some features of *correlation* between the underlying utility functions. Traditionally, instead, the central case of RUM has been that of *independent* distributions (see Tversky [15] and the discrete choice literature). To illustrate, recall that for independent RUM the following *order* property must hold:

$$p(x, xy) = 1 = p(z, zw) \Rightarrow p(x, xw) = 1 \text{ or } p(z, zy) = 1$$

To see that this property does not necessarily hold for random checklists with a menu-independent π, consider the following

Example 13

	P	Q	R
x	1	0	1
y	0	0	1
z	1	1	0
w	0	1	0

with $\pi(PQR) = \pi(PRQ) = \pi(QPR) = \pi(RPQ) = \frac{1}{4}$*. We have* $p(x, xy) = 1 = p(z, zw)$ *and yet* $p(w, wx) = \pi(QPR) = \frac{1}{4}$ *and* $p(y, zy) = \pi(RPQ) = \frac{1}{4}$.

This example can be interpreted along the lines of the classical Rome/Paris example used against the independent RUM model (see e.g. Tversky [15]). Interpret x and y as holiday packages to Rome and z and w as holiday packages to Paris. However, x and z also incorporate an extra bonus. The example envisages the agent being genuinely undecided between Rome and Paris, but always preferring having the bonus for a given destination. The random checklist described has the agent being in the mood for Rome or for Paris with even chances; property Q details the special attractions of Rome and property R those of Paris, while property P is 'having a special bonus'.

The moral of this example is that while a CRC with menu-independent π is equivalent to a RUM, the differences between the primitives of the two models are important: some assumptions that are focal in one model (such as independence for RUM) are not at all so in the other. Indeed, one of the reasons for the origin of EBA and the consideration of properties in the cognitive mechanism underlying choice was *precisely* to explain Rome-Paris sort of examples.

The above example also shows that a choice by random checklist does not necessarily reveal any stochastic preference (a terminology introduced in [4]). This happens when $p(x, x \cup A) \geq p(y, y \cup A) \Rightarrow p(x, x \cup B) \geq p(y, y \cup B)$ whenever $\{x, y\} \cap (A \cup B) = \varnothing$. That is, if x is chosen with a higher probability than y over a set of alternatives, this should remain true independently of the set of alternatives other than x and y. In the example, we have $p(x, xy) = 1 > \frac{3}{4} = p(z, zy)$ but $p(x, xw) = \frac{3}{4} < 1 = p(z, zw)$, so that no stochastic preference is revealed.

Finally, we have not shown explicitly that CRC implies EBA, but this is a consequence of the fact that CRC extends RUM and RUM extends EBA (Tversky [16]). So we can state:

Corollary 14 *A random choice p that is an EBA is also a CRC. In addition, the CRC can be chosen with a menu-independent π.*

6 Conclusion

In this paper we have clarified some connections, as well as points of difference, between psychologically based sequential decision heuristics and random utility maximisation (RUM). We have arrived at strong and behaviourally interpretable testable restrictions for our proposed random checklist model, which generalises both RUM and Elimination by Aspects.

We have fallen short of a full characterisation, and we only have an algorithmic (linear system) characterisation for a special case. It would, of course, be a desirable objective to identify completely a set of testable and behaviourally interpretable implications of the theory. But, in view of the fact that an equivalent full characterisation for the RUM model has escaped theorists for over fifty years, this may be a tall order.

The random checklist model, while equivalent in a special case to the RUM model, incorporates more primitives and thus a richer source of interesting potential identification problems. For example, we may conceive of situations where an external observer of the choice probabilities is also informed of a subset of the properties in a subset of moods, if it were the case that the agent is using a random checklist. The question is whether the choice probabilities can be induced by a random checklist that is consistent with the prior information of the observer. This question, which cannot be posed in the RUM framework, is a subject for future research.

References

[1] Barbera, S. and P. Pattanaik (1986) "Falmagne and the Rationalizability of Stochastic Choice in terms of Random Orderings," *Econometrica* 54: 707-15.

[2] Block, H.D. and J. Marschak (1960) "Random Orderings and Stochastic Theories of Responses", in Olkin, Ingram, Sudish G. Gurye, Wassily Hoeffding, Willima G. Madow, and Henry B. Mann (eds.) Contributions to Probability and Statistics, Stanford University Press, Stanford, CA.

[3] Gul, F. and W. Pesendorfer (2006) "Random Expected Utility", *Econometrica* 74: 121-146.

[4] Gul, F., P. Natenzon and W. Pesendorfer (2010) "Random Choice as Behavioral Optimization", mimeo, Princeton University.

[5] Luce, R.D., (1959) *Individual choice behavior: a theoretical analysis*, Wiley, New York.

[6] Marschak, J. (1960) "Binary choice constraints and random utility indicators", Cowles Foundation Paper 155.

[7] Mandler, M., P. Manzini and M. Mariotti (2012) "One Million Answers to Twenty Questions: Choosing by Checklist" *Journal of Economic Theory* 147: 71-92

[8] McFadden, D.L. (1974) "Conditional Logit Analysis of Qualitative Choice Behavior", in P. Zarembka (ed.), Frontiers in Econometrics, p. 105-142, Academic Press: New York.

[9] McFadden, D.L. (1974) "The Measurement of Urban Travel demand", *Journal of Public Economics* 3: 303-28.

[10] McFadden, D.L. (2000) "Economic choices", *American Economic Review*, 91 (3): 351-378.

[11] McFadden, D.L. (2005) "Revealed Stochastic Preference: A Synthesis", *Economic Theory* 26: 245-64.

[12] McFadden, D.L. and K. Richter (1991) "Stochastic Rationality and Revealed Stochastic Preference", in J. Chipman, D. McFadden, K. Richter (eds) Preferences, Uncertainty, and Rationality, Westview Press, 161-186.

[13] Richter, M.K. (1966) "Revealed Preference Theory", *Econometrica* 34(3): 635-45.

[14] Sattath, S. and A. Tversky (1976) "Unite and conquer: A multiplicative inequality for choice probabilities", *Econometrica*, 44: 78–89.

[15] Tversky, A. (1972) "Choice by Elimination", *Journal of Mathematical Psychology* 9: 341-67.

[16] Tversky, A. (1972) "Elimination by Aspects: A Theory of Choice", *Psychological Review* 79: 281-99.